

J02275. Minute virus of m...[gi:332293]

LOCUS MVMPG 5149 bp ss-DNA linear VRL 22-MAY-1995

DEFINITION Minute virus of mice, complete genome.

ACCESSION J02275 M12520 M12521 M14704

VERSION J02275.1 GI:332293

KEYWORDS alternative splicing; capsid protein; complete genome;
nonstructural protein.

SOURCE Mice minute virus

ORGANISM Mice minute virus

Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

REFERENCE 1 (bases 1 to 5149)

AUTHORS Astell,C.R., Thomson,M., Merchlinsky,M. and Ward,D.C.

TITLE The complete DNA sequence of minute virus of mice, an autonomous
parvovirus

JOURNAL Nucleic Acids Res. 11 (4), 999-1018 (1983)

MEDLINE 83143341

PUBMED 6298737

REFERENCE 2 (bases 1 to 5149)

AUTHORS Astell,C.R., Gardiner,E.M. and Tattersall,P.

TITLE DNA sequence of the lymphotropic variant of minute virus of mice,
MVM(i), and comparison with the DNA sequence of the fibrotropic
prototype strain

JOURNAL J. Virol. 57 (2), 656-669 (1986)

MEDLINE 86115415

PUBMED 3502703

REFERENCE 3 (sites)

AUTHORS Morgan,W.R. and Ward,D.C.

TITLE Three splicing patterns are used to excise the small intron common
to all minute virus of mice RNAs

JOURNAL J. Virol. 60 (3), 1170-1174 (1986)

MEDLINE 87061199

PUBMED 3783817

COMMENT Original source text: Minute virus of mice (strain MVM(p)), passed
in mouse l (variant A-9) cells.

The parvoviridae family contains two groups that infect mammalian
hosts: (i) defective (helper-dependent) adeno-associated viruses,
and (ii) autonomous (helper-independent) parvoviruses. MVM is a
member of the latter group. Both groups have been demonstrated to
package both plus and minus strands (in separate particles) of the
ss-DNA genome, though the minus strand is more typically packaged
in the latter group.

The sequence below corresponds to the plus (+) strand, also

referred to as the C-strand. The minus (-) strand is also referred to as the V-strand.

The 3' and 5' termini both exhibit the potential for forming stable 'fold-back' hairpins; these sequences appear to play a role in replication [1].

revision 4804 4870 a-65bp-a in [2]; aa in [1] [2]

revises [1].

ORIGIN 5' end of genome; 415 bp upstream of PstI site.

```

1 atttttagaa ctgaccaacc atgttcacgt aagtacgtg atgacgcgcg ctgcgcgcgc
61 gccttcggac gtcacacgtc acttacgttt cacatgggtg gtcagttcta aaaatgataa
121 gcggttcagg gagtttaaac caaggcgcga aaaggaagtg ggcgtgggtt aaagtatata
181 agcaactact gaagtcagtt acttatcttt tctttcattc tgtgagtcga gacgcacaga
241 aagagagtaa ccaactaacc atggctggaa atgcttactc tgatgaagtt ttgggagcaa
301 ccaactgggt aaaggaaaaa agtaaccagg aagtgttctc atttgtttt aaaatgaaa
361 atgttcaact gaatggaaaa gatatcggat ggaatagtta caaaaagag ctgcaggagg
421 acgagctgaa atctttacaa cgaggagcgg aaactacttg ggaccaaagc gaggacatgg
481 aatgggaaac cacagtggat gaaatgacca aaaagcaagt attcattttt gattctttgg
541 ttaaaaaatg tttatttgaa gtgcttaaca caaagaatat atttctggt gatgttaatt
601 ggtttgtgca acatgaatgg ggaaaagacc aaggctggca ctgccatgta ctaattggag
661 gaaaggactt tagtcaagct caagggaat ggtggagaag gcaactaaat gtttactgga
721 gcagatgggt ggtaacagcc tgtaatgtgc aactaacacc agctgaaaga attaaactaa
781 gagaaatagc agaagacaat gagtgggtta ctctacttac ttataagcat aagcaaacca
841 aaaaagacta taccaagtgt gttcttttg gaaacatgat tgcttactat ttttaacta
901 aaaagaaaat aagcactagt ccaccaagag acggaggcta tttcttagc agtgactctg
961 gctggaaaac taacttttta aaagaaggcg agcgccatct agtgagcaa ctatacactg
1021 atgacatgcg gccagaaacg gttgaaacca cagtaaccac tgcgcaggaa actaagcgcg
1081 gcagaattca aactaaaaaa gaagtttcta ttaaaactac acttaaagag ctggtgcata
1141 aaagagtaac ctaccagag gactggatga tgatgcagcc agacagttac attgaaatga
1201 tggctcaacc aggtggagaa aacctgctga aaaatacgct agagatttgt aactaactc
1261 tagccagaac caaaacagca ttgacttaa ttttagaaaa agctgaaacc agcaaactaa
1321 ccaacttttc actgcctgac acaagaacct gcagaatttt tgcttttcat ggctggaact
1381 atgttaaagt ttgccatgct atttgcgtg ttttaaacag acaaggaggc aaaagaaata
1441 ctgttttatt tcatggacca gccagcacag gcaaacttat tattgcacaa gccatagcac
1501 aagcagttgg caatgttggt tgctataatg cagccaatgt aaactttcca ttaatgact
1561 gtaccaacaa gaacttgatt tgggtagaag aagctggtaa ctttggacag caagtaaacc
1621 agtttaaagc catttgctct ggtcaacta ttcgcattga tcaaaaagga aaaggcagca
1681 aacagattga accaacacca gtcatcatga ccacaaatga gaacattaca gtggtcagaa
1741 taggctgcga agaaagacca gaacacactc aaccaatcag agacagaatg cttaacattc
1801 atctaacaca taccttgctt ggtgactttg gtttgggtga caaaaatgaa tggcccatga
1861 tttgtgcttg gttggtaaag aatggttacc aatctaccat ggcaagctac tgtgctaaat
1921 ggggcaaaag tcttgattgg tcagaaaact gggcggagcc aaagtgcca actcctataa
1981 atttactagg ttcggcacgc tcaccattca cgacaccgaa aagtagcct ctacgccaga
2041 actatgcact aactccactt gcatcggatc tcgaggacct ggcttagag ccttggagca

```

2101 caccaaatac tcctgttgcg ggcactgcag aaaccagaa cactggggaa gctggtcca
 2161 aagcctgcc aagatgtcaa ctgagcccaa ctgtgtcaga gatcaggag gatttgagag
 2221 cgtgcttcgg tgcggaaccg ttgaagaaag acttcagcga gccgctgaac ttggactaag
 2281 gtacgatggc gcctccagct aaaagagcta aaagaggtaa gggtttaagg gatggttggt
 2341 tgggtgggta ttaatgttta attacctgtt ttacaggcct gaaatcactt ggttttaggt
 2401 tgggtgcctc ctggctacaa gtacctggga ccagggaaca gccttgacca aggagaacca
 2461 accaatccat ctgacgccgc tgccaaagag cacgacgagg cctatgatca atacatcaaa
 2521 tctggaaaaa atccttacct gtacttctct gctgctgac aacgctttat tgaccaaacc
 2581 aaggacgcca aagactgggg aggcaagggt ggtcactact ttttagaac caagcgcgct
 2641 tttgcaccta agcttgctac tgactctgaa cctggaactt ctggtgtaag cagagctggt
 2701 aaacgacta gaccactgc ttacattttt attaaccaag ccagagctaa aaaaaactt
 2761 acttcttctg ctgcacagca aagcagtcaa accatgagt atggcaccag ccaacctgac
 2821 agcggaaacg ctgtccactc agctgcaaga gttgaacgag cagctgacgg cctggaggc
 2881 tctgggggtg ggggctctgg cgggggtggg gttggtgtt ctactgggtc ttatgataat
 2941 caaacgcatt atagattctt gggtagcggc tggtagaaa ttactgact agcaactaga
 3001 ctagtacatt taaatagcc taaatcagaa aactattgca gaatcagagt tcacaatata
 3061 acagacacat cagtcaaagg caacatggca aaagatgatg ctcatgagca aatttgaca
 3121 ccatggagct tgggtgatgc taatgcttgg ggagtttggc tccagccaag tgactggcaa
 3181 tacattgca acacatgag ccagcttaac ttggtatcac ttgatcaaga aatattcaat
 3241 gtagtgctga aaactgttac agagcaagac ttaggaggtc aagctataaa aatatacaac
 3301 aatgacctta cagcttgcac gatggttgca gtagactcaa acaacatttt gccatacaca
 3361 cctgcagcaa actcaatgga aacacttggg ttctaccctt ggaaaccaac catagcatca
 3421 ccatacaggt actatttttg cgttgacaga gatctttcag tgacctacga aatcaagaa
 3481 ggcacagttg aacataatgt gatgggaaca ccaaaggaa tgaattctca atttttacc
 3541 attgagaaca cacaacaaat cacattgctc agaacagggg acgaatttgc cacaggtact
 3601 tactactttg acacaaattc agttaaactc acacacacgt ggcaaaccaa ccgtcaactt
 3661 ggacagcctc cactgctgct aacctttcct gaagctgaca ctgatgcagg tacacttact
 3721 gctcaaggga gcagacatgg aacaacaaa atgggggtta actgggtgag tgaagcaatc
 3781 agaaccagac ctgctcaagt aggattttgt caaccacaca atgactttga agccagcaga
 3841 gctggaccat ttgctgccc aaaagtcca gcagatatta ctcaaggagt agacaaagaa
 3901 gccaatggca gtgttagata cagttatggc aaacagcatg gtgaaaattg ggcttccat
 3961 ggaccagcac cagagcgcta cacatgggat gaaacaagct ttggttcagg tagagacacc
 4021 aaagatgggt ttattcaatc agcaccacta gttgttcac caccactaaa tggcattctt
 4081 acaaatgcaa accctatttg gactaaaaat gacattcatt ttcaaatgt ttttaacagc
 4141 tatggtccac taactgcatt ttacaccca agtctgtat accctcaagg acaaatatgg
 4201 gacaaagaac tagatcttga acacaaacct agacttaca taactgctcc attgtttgt
 4261 aaaaacaatg cacctggaca aatgttgggt agattaggac caaacctaac tgaccaatat
 4321 gatccaaacg gagccacact ttctagaatt gttacatacg gtacatttt ctggaaagga
 4381 aaactaacca tgagagcaaa acttagagct aacaccactt ggaaccaggt gtaccaagta
 4441 agtgctgaag acaatggcaa ctacataatg agtgtaacta aatgggttacc aactgctact
 4501 ggaaacatgc agtctgtgcc gcttataaca agacctgttg ctagaaatac ttactaacta
 4561 accatgcttt ttcttctgt acttcatata ttattaagac taataaagat acaacataga
 4621 aatataatat tacgtataga tttaagaaat agaataatat ggtacttagt aactgttaa

4681 aataatagaa cctttggaat aacaagatag ttagttgggt aatgtagat agaataagaa
4741 gatcatgtat aatgaataaa aggggtggaag ggtgggttgg aggttaatgt tagatagaat
4801 aagaagatca tgtataatga ataaaagggt ggaagggtgg ttggtaggta ttccttaga
4861 ctgatgtta aggacaaaa aaataataaa acttttttaa aactcaacca agactactgt
4921 ctattcagtg aaccaactga accattagta ttactatgtt ttaggggtgg gaggggtggga
4981 gatacatgtg ttcgctatga gcgaactgggt actgggttgg tgctctgctc aaccaaccag
5041 accggcaaag ccggtctggt tggttgagcg caaccaacca gtaccagttc gtcatagcg
5101 aacacatgta tctcccaccc tcccacccta aaaacatagt aataactaat

NC_004713. LuIII virus, comp...[gi:29742044]

LOCUS NC_004713 5135 bp ss-DNA linear VRL 20-AUG-2003

DEFINITION LuIII virus, complete genome.

ACCESSION NC_004713

VERSION NC_004713.1 GI:29742044

KEYWORDS

SOURCE LuIII virus (LuIIIV)

ORGANISM LuIII virus

Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

REFERENCE 1 (bases 1 to 5135)

AUTHORS Diffoot,N., Chen,K.C., Bates,R.C. and Lederman,M.

TITLE The complete nucleotide sequence of parvovirus LuIII and localization of a unique sequence possibly responsible for its encapsidation pattern

JOURNAL Virology 192 (1), 339-345 (1993)

MEDLINE 93297126

PUBMED 8517025

COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence was derived from M81888.

Coding regions were annotated at the NCBI following the annotation of closely related Mouse parvovirus 1 (U12469).

```

1 atcattttta gaactaacca accatgttca cgtaagtac gtgatgacgc gcgctacgcg
61 cgctgccttc ggcagtcaca cgtcacttac gtctcacatg gttggtagt tctaaaaatg
121 ataagcgggt cagggagttt aaaccaaggc gcgaaaagga agtgggcgtg gttttaagta
181 tataagcgac acgttaagtc agttacttac tcttcgctt attctgtaag tcgagacaca
241 cagagtaacc aactaacca ctagccatgg ctggaaacgc gtactctgat gaagttttgg
301 gaacaactaa ctgggtgaag gataagagca accaggaagt attctcattt gtttttaaaa
361 atgaggatgt tcagctcaat ggaaaaata tcggatggaa cagttacaga aaggagctgc
421 aagaggagga gctgaaatct ttacaacgag gagctgaaac tacctgggac cagagcgagg
481 acatggaatg ggaatcttca gtggatgaac tgaccaaaaa gcaagtattc attttgact
541 ctttagttaa aaagtgtctc ttgaagtac tgagcacaaa gaacatagct ctagtgatg
601 ttacttggtt tgtacagcat gaatggggaa aagaccaagg ctggcactgt catgtgctca
661 ttggaggcaa gaactttagc caggctcaag gaaaatggtg gaggagacaa ttaaatttt
721 actggagtag atggttggtg acagcctgta gcgtgcagct atcaccagct gaaagaatta
781 aactaagaga aatagcagaa gaccaagaat gggttactct gcttacttat aagcataagc
841 aaacaaaaaa agactatact aagtgtgttt gctttggaaa tatggttgct tactactttt
901 taacaaaaaa gaaaatatgt accagtccac caaggacggg aggctatttt cttagtagt
961 actctggctg gaaaactaac ttttgaaag aaggcgaacg ccatctagtg agcaaactat
1021 atactgatga catgcggcca gaaacggtg agaccacagt aaccacagcg caggaaacta
1081 agcgcggcag aattcaactt aagaaggaag tctctattaa gactacactt aaagagctgg
1141 tacataagag agtaacctca ccagaagact ggatgatgat gcagccagac agttacattg

```

1201 aaatgatggc tcaaccaggg ggagaaaacc tacttaagaa tacgctagag atctgtacgc
 1261 tgactctagc cagaacaaaa acagcctttg acttgatttt agaaaaagct gaaaccagca
 1321 aactaacaa cttttactg gctgatacaa gaacctgtag aatctttgct tttcatggct
 1381 ggaactacat caaagtctgt catgctatgt gttgtgtctt gaacagacag ggaggcaaaa
 1441 gaaatactgt tctgtttcat ggaccagcca gtacaggcaa atcaatcatt gcacaggcca
 1501 tagcacaggc agttggtaat gttggttgtt ataacgcagc caatgtgaac ttccattta
 1561 atgactgtac caacaagaac ttaatctggg tggaagaagc tggttaacttt ggacagcaag
 1621 taaaccagtt taaagccatt tgttctggtc agaccattcg cattgaccaa aaaggaaaag
 1681 gcagcaaaaca gattgaacca acaccagtga tcatgaccac aaatgaaaac atcacagtgg
 1741 tcaaaatagg gtgtgaagag agaccagaac aactcaacc aatcagagac agaattgtaa
 1801 acattcatct gacacataca ttgcctgggtg actttgggtt ggttgataaa aacgaatggc
 1861 ctatgatatg tgcttgggtg gtaagaacg gttaccaatc gaccatggca agttactgtg
 1921 ctaaattgggg caaagttcct gattggacag aaaactgggc ggagccaaaa gtaacgactg
 1981 aaataaatc ggtaggttca accaactcac catctccgaa aagtacgct ctcagccaga
 2041 actacgcact aactccgtcg gatctcgagg acctggctct ggagccttgg agcacaccaa
 2101 gtactcctgt tgtgggcact gtcaaaacc cgaacactgg ggaaactggt tcaacagcct
 2161 gtcaagaagc tcaacggagc ccaacttggc cgagatcga ggaggattg agagcgtgtg
 2221 tcagttcgga aactggaag agcgactccg aacagctacc aaacttggat taaggtagca
 2281 tggcgctcc ggctaaaaga gctaaaagag gtaaggggtt aagggatggt tggtaggttg
 2341 gtgggttatt aatatgtgac tacctgtttt acaggcctga aatcacttgg ttctaggttg
 2401 ggtgcctcca ggctacaagt acctgggacc agggaacagc ctaaccaag gagaaccaac
 2461 caatccatct gacgtgtg ctaaagagca cgacgaggcc tacgaccaat acatcaaatc
 2521 tggaagaat ccttacctgt acttctctcc tgctgatcaa cgcttcattg accaaaccaa
 2581 agacgctaaa gactggggcg gaaaggttgg tcactactt ttagaacca agcgtgcttt
 2641 tgcacctag ctttctactg actctgagcc tgggacttct ggtgtgagca cagctggtaa
 2701 acgtactaaa ccacctgctc acatctttat taaccaagcc agggctaaaa aaaacgtac
 2761 ttctcttct ggcgagcaga ggactcagac aatgagtgt ggcaccgacc aatctgacag
 2821 cggaaacgct gtccagttag ctgctagagt tgagcgagca gctgacggc ctggaggctc
 2881 tggggggcggg ggctctggtg ggggtgggtt tggcgttct actggcagtt atgataatca
 2941 aacacattat aagtttctag gggatgggtg gtagagatt actgcttaca gcacacgat
 3001 ggtacacttg aacatgccta aatcagaaaa ctacttagg gtgcgcgtac acaacacaaa
 3061 tgacacaggt acagcaagtc acatggctat ggacgatgct catgaacaga ttggacacc
 3121 atggagtctg gttgatgcta atgcttgggg agtttgggtt caaccaagt actggcagta
 3181 catttcta ataatgattc acatcaattt acattcactt gaccaagaat tgtttaatgt
 3241 ggtcatcaaaa acagtgtg aacagaacac aggagctgag gccattaagg tctacaacaa
 3301 tgacctcact gctgcatga tggttgctct tgattctaac aacatactgc cttacacacc
 3361 agccatagac aatcaagaga cacttgggtt ctatccatgg aaaccaacca taccaagtcc
 3421 ttacagatac tattttagct gtgacagaaa cttatcagtt acttacaag acgaagcagg
 3481 aaccatcact gacacaatgg gtttggccag tggcctgaac tccaatttt ttaccattga
 3541 gaacactcag cgtattaacc tactcagaac tggggatgag tatgctactg gaacttacta
 3601 ctttgacaca gaaccaatca gactaactca cacgtggcaa accaacagac acctgggtca
 3661 gcctccacaa attactgaac taccaagctc tgacactgct aacgctactt taacagctag
 3721 aggttacaga tcaggtctga ctcaaatca aggcagaaat gatgtgactg aagctactag

3781 ggtcagacct gcacaggttg gattttgtca gcctcatgac aattttgaaa ccagcagagc
3841 ggggcctttc aagggtccgg tagtgccagc agacatcaca caaggcctag accatgatgc
3901 caatggtagc ctgagatata cctatgacaa acaacatggt caaagctggg caagtcagaa
3961 caacaaagac aggtacactt gggatgctgt taactatgat tctggcagat ggactaaca
4021 ctgttttatt caatcagtag catttacatc agaaccaaat gctaaccaaa tacttactaa
4081 ccgtgacaac ctagcgggta agactgacat acattttacc aacgcattta acagttatgg
4141 accactaact gcttttccac atcctgcgcc gatttaccca caagggcaga ttgggacaa
4201 agaacttgat ctgaaacaca agccaagact gcacacacag gtccttttg tctgtaaaaa
4261 caatgctcca ggtagcttc tggtaggct agcacctaac ttgactgacc agtatgatcc
4321 taatagtctt aacctatcta gaattgtcac ctatggcacc ttcttctgga agggcaaact
4381 aactctaaaa gcaaagatga gacctaatgc tacttggaac ccagtcttcc aaataagtgc
4441 taccaacca ggaaccaatg actacatgag cattgaaaga tggttacca ctgctactgg
4501 caacataaca aatgtgcctc tgctttctag acctgttgct agaaacactt actaactaac
4561 tatgctctat gcttcatata tattatata attattatac taactaacca tgtttactct
4621 tacattactt catataatat taagactaat aaaatacaa catagaaata taatattaca
4681 tatagatata agaatagaa taatatggta ctacttact gtagaaata atagaacttt
4741 tggataaca agatagttag ttggtttatg ttatatagaa tataagaaga tgatgtacaa
4801 agaataaaaag ggtgggaggg tgggtgggtg gtactccctt agactgaatg ttagggacca
4861 aaaaaataat aaaattcttg aaaaccaac aaggactact gtcatttca gtgaaccaac
4921 tgaaccatta gtatcaatat gattttaggg tgggggggtg ggagatacat atgttacta
4981 tggaccaact ggtactggtt ggttgctctg ctccaacca ccagaccggc tctgccggtc
5041 tggttgggtg agcgcaacca accagtacca gttgtccat agtgaacata tgtatctccc
5101 accccccac ctaaaaaca tattgatact aatgg

1: NC_001358. Parvovirus H1, co...[gi:9626078]

Links

LOCUS NC_001358 5176 bp ss-DNA linear VRL 20-AUG-2003

DEFINITION Parvovirus H1, complete genome.

ACCESSION NC_001358

VERSION NC_001358.1 GI:9626078

KEYWORDS genome; origin of replication.

SOURCE Parvovirus H1

ORGANISM Parvovirus H1

Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.

REFERENCE 1 (bases 1 to 4534)

AUTHORS Rhode,S.L. III and Paradiso,P.R.

TITLE Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid-arrested translation

JOURNAL J. Virol. 45 (1), 173-184 (1983)

MEDLINE 83112183

PUBMED 6823009

REFERENCE 2 (bases 4435 to 5176)

AUTHORS Rhode,S.L. III and Klaassen,B.

TITLE DNA sequence of the 5' terminus containing the replication origin of parvovirus replicative form DNA

JOURNAL J. Virol. 41 (3), 990-999 (1982)

MEDLINE 82242308

PUBMED 6284985

COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence was derived from X01457.

The viral genome (- strand) is the complementary strand to that shown below (+ strand).

[1] discusses other major open reading frames, but was uncertain as to exact boundaries and/or splicing locations. the non-capsid protein in the features table is speculatively identified as the rf rep gene product: either the postulated site-specific nickase, or the terminal bound protein, or both [1].

ORIGIN

```

1 catttttaga actgaccaac catgttcacg caagtgcgt gatgacgcgc gctgcgcgcg
61 ctgccttcgg cagtacacg tctactagct ttacatggt tggtcagttc taaaaatgat
121 aagcggttca gagagtttga aaccaaggcg ggaaacggaa gtgggcgtgg ctaactgtat
181 ataagcagtc actctggctg gttactcact ctgctttcat ttctgagttt gtgagacaca
241 ggagcgagac taaccaacta accatggctg gaaacgctta ctccgatgag gttttgggag
301 taacaaactg gctgaaggac aaaagtagcc aggaggtgtt ctatttgtt tttaaaatg

```


361 aaaacgtcca actaaatgga aaggacatcg gttggaatag ttacagaaag gagctacaag
 421 atgacgagct gaagtctcta caacgagggg cgagagaccac ttgggaccaa agcgaggaca
 481 tggaatggga gagcgagctg gatgacatga ccaaaaagca agtatttatt ttgattctt
 541 tggtaagaa gtgtttgtt gaagtgtca gcacaaagaa catagctcct agtaatgta
 601 cttggttcgt gcagcatgaa tggggaaagg acccaggctg gcactgtcat gtgctgattg
 661 gaggcaagga ctttagtcaa cctcaaggaa aatggtggag aaggcagcta aatgtgtact
 721 ggagtagatg gttggtgact gcctgtaatg ttcaactaac accagctgaa agaattaaac
 781 tgagagaaat agcagaggac agtgaatggg tcaacttgcct tacctataag cataagcaca
 841 ccaagaagga ctataccaag tgtgttctt ttggaacat gattgcttat tacttttaa
 901 gcaaaaagaa aatatgtacc agtccaccaa gggacggagg ctatttctt agcagtgact
 961 ctggctggaa aactaactt ttgaaagagg gcgagcgcca tctagtgagc aaactgtata
 1021 ctgatgagat gaaaccagaa acggtcgaga ccacagtac cactgcacag gaagctaagc
 1081 gcggcagaat tcaactaga gaggaggtct cgattaaaac cacactcaa gagttggtac
 1141 ataaaagagt aacctacca gaagactgga tgatgatgca gccagacagt tacattgaaa
 1201 tgatggctca accaggtgga gaaaacttgc ttaaaatac actagagatc tgtactga
 1261 ctctagcaag aaccaaaca gccttgact tgattctgga aaaagctgaa accagcaaac
 1321 tagccaactt ttccatggct agcaccagaa cctgtagaat cttgctgag catggctgga
 1381 actatattaa agtctgccat gccatctgt gtgtgctgaa tagacaagga ggcaaaagga
 1441 acactgtgct ctttcacgga ccagccagca caggcaaac tattattgca caagccatag
 1501 cacaagcagt tggtaatgtt ggtgttaca atgctgcaa tgtgaactt ccattaatg
 1561 actgtaccaa caaaaactg atttgggtg aagaagctgg taactttggc cagcaagtaa
 1621 accaattcaa agctattgt tctggccaa ccatacgcat tgatcaaaa ggaaaaggca
 1681 gcaaacagat tgaaccaaca ccagttatta tgaccacaa cgagaacatt accgtggtta
 1741 gaataggctg tgaggaaaga ccagaacaca ctcaaccaat cagagacaga atgtcaaca
 1801 ttcacgtgac acgtacacta cctggtgact ttggttggg ggataagcac gaatggcctc
 1861 tgatctgtgc ttggttggg aagaatggtt accaatctac catggctgt tactgtgcta
 1921 aatggggcaa agttctgat tggcagagg actggcgga gccgaagcta gacactccta
 1981 taaattcgt aggttcaatg cgctcaccat ctctgactcc gagaagtacg cctctagcc
 2041 aaaactacgc tcttactcca cttgcatcgg acctgcgga cctagctcta gagccttga
 2101 gcacacaaa tactcctgtt gcgggactg cagcaagcca aaactggg gaggctggtt
 2161 ccacagcctg ccaaggtgct caacggagcc caacctggc cgagatcgag gcggattga
 2221 gagcttgctt cagtcaagaa cagtggaga gcgactcaa cgaggagctg acctggact
 2281 aaggtacaat ggcacctcca gctaaaagag ctaaaagagg taaggggcta agggatggtt
 2341 ggttgggtgg gtactaatgt atgactacct gtttacagg cctgaaatca cttggttcta
 2401 ggttgggtgc ctctggcta caagtacct ggaccaggga acagcctga ccaaggagaa
 2461 ccaaccaacc ctctgacgc cgctgcaaaa gaacacgacg aagcctacga ccaatacatc
 2521 aaatctggaa aaaatcctta cctgtacttc tctctgctg atcaacgctt cattgacaa
 2581 accaaagacg ccaaggactg gggcggaag gttggtcact acttttttag aaccaagcga
 2641 gctttgca ctaagcttc tactgactct gaacctggca cttctggtg gagcagacct
 2701 ggtaaacgaa ctaaaccacc tgctcacatt tttgaaac aagccagagc taaaaaaaaa
 2761 cgcgttctc ttgctgcaca gcagaggact ctgacaatga gtgatggcac cgaaacaac
 2821 caaccagaca ctggaatcgc taatgctaga gttgagcgat cagctgacgg aggtggaagc
 2881 tctgggggtg ggggctctgg cgggggtggg attggtgtt ctactgggac ttatgataat

2941 caaacgactt ataagttttt gggagatgga tgggtagaaa taactgcaca tgcttctaga
 3001 cttttgactt tgggaatgcc tccttcagaa aactactgcc gcgtcaccgt tcacaataat ,
 3061 caaacaacag gacacggaac taaggtaaag ggaaacatgg cctatgacac acatcaacaa
 3121 atttggacac catggagctt ggtagatgct aatgcttggg gagtttggtt ccaaccaagt
 3181 gactggcagt tcattcaaaa cagcatggaa tcgctgaatc ttgactcatt gagccaagaa
 3241 ctatttaatg tagtagtcaa aacagtcact gaacaacaag gagctggcca agatgccatt
 3301 aaagtctata ataatgactt gacggcctgt atgatggttg ctctggatag taacaacata
 3361 ctgccttaca cacctgcagc tcaaactca gaaacacttg gtttctacc atggaacaa
 3421 accgcaccag ctccttacag atactacttt ttcatgccta gacaactcag tgtaacctt
 3481 agcaactctg ctgaaggaac tcaaatcaca gacaccattg gagagccaca ggcactaaac
 3541 tctcaatttt ttactattga gaacaccttg cctattactc tctgcgcac aggtgatgag
 3601 tttaactg gcacctacat cttaacact gaccactta aacttactca cacatggcaa
 3661 accaacagac acttggcatg cctccaagga ataactgacc taccaacatc agatacagca
 3721 acagcatcac taactgcaa tggagacaga ttggatcaa cacaacaca gaatgtgaac
 3781 tatgtcacag aggtttgcg caccaggcct gctcagattg gcttcatgca acctcatgac
 3841 aactttgaag caaacagagg tggccattt aaggttcag tggtagcgt agacataaca
 3901 gctggcgagg accatgatgc aaacggagcc atacgattta actatggcaa acaacatggc
 3961 gaagattggg ccaaacaagg agcagcacca gaaaggtaca catgggatgc aattgatagt
 4021 gcagctggga gggacacagc tagatgctt gtacaaagt caccaatc tatteacca
 4081 aacaaaacc agatcttga gcgagaagac gccatagctg gcagaactaa catgcattat
 4141 actaatgtt ttaacagcta tggccactt agtgcattc ctcattcaga tccatttat
 4201 ccaatggac aaatttggga caaagaattg gacctggaac acaaacctag actacagta
 4261 actgcacat ttgtttgta aaacaacca ccaggtaac tattgttca ctggggcct
 4321 aatctgactg accaatttga ccaaacagc acaactgtt ctcgactgt tacatatagc
 4381 acttttact ggaagggtat ttgaaattc aaagccaaac taagaccaa tctgacctg
 4441 aatctgtat accaagcaac cacagactt gttgccaatt ctacatgaa tgttaagaa
 4501 tggctcccat ctgcaactgg caacatgcac tctgatccat tgattttag acctgtcct
 4561 cacatgacat actaaccaac caactatgt tctctgttg cttcacataa tacttaact
 4621 aactagacta caacataaaa atatacactt aataatagat tattaataat aacataatat
 4681 ggtagggtta ctgtaaaaa taatagaact ttggaataa atatagttag ttggttaatg
 4741 ttagatagaa tataaaaaga tttgtattt taaaataaat atagttagt ggtaatgtt
 4801 agatagaata taaaagatt ttgtatttg gaaataaaa ggggtggttg gtggttggt
 4861 ggtactcct tagactgaat gttagggacc aaaaaataa taaaataatt aaatgaaca
 4921 aggactactg tctattcagt tgaccaactg aacctatagt atcactatgt ttttaggtg
 4981 ggggggtgg agatacatac gttcgctatg gaccaagtgg taccggttg ttgctaagct
 5041 cgaacaagac ggctaagccg gtccggttg ttgagcgcaa ccaaccgga ccacttggtc
 5101 catagcgaac gtatgtatc cccaccccc caccataaa acatagtat actataggt
 5161 cagttgtca actgaa